

RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.

Application Serial Number: 09/930,591
Source: IFN16
Date Processed by STIC: 06-08-2005

ENTERED

RE-RUN.



IFW16

RAW SEQUENCE LISTING

DATE: 06/08/2005

PATENT APPLICATION: US/09/930,591

TIME: 10:52:17

Input Set : N:\Crf4\Refhold\09_folder\I930591.raw

Output Set: N:\CRF4\06082005\I930591.raw

```

1 <110> APPLICANT: Matti Sallberg
2 <120> TITLE OF INVENTION: A HEPATITIS C VIRUS NON-STRUCTURAL
3   NS3/4A FUSION GENE
4 <130> FILE REFERENCE: TRIPEP.028AUS
5 <140> CURRENT APPLICATION NUMBER: US/09/930,591
6 <141> CURRENT FILING DATE: 2001-08-15
7 <150> PRIOR APPLICATION NUMBER: 60/225,767
8 <151> PRIOR FILING DATE: 2000-08-17
9 <150> PRIOR APPLICATION NUMBER: 60/229,175
10 <151> PRIOR FILING DATE: 2000-08-29
11 <150> PRIOR APPLICATION NUMBER: 09/705,547
12 <151> PRIOR FILING DATE: 2000-11-03
13 <160> NUMBER OF SEQ ID NOS: 34
14 <170> SOFTWARE: FastSEQ for Windows Version 4.0
16 <210> SEQ ID NO: 1
17 <211> LENGTH: 2061
18 <212> TYPE: DNA
19 <213> ORGANISM: Artificial Sequence
20 <220> FEATURE:
21 <223> OTHER INFORMATION: Hepatitis C virus NS3/4A coding region
22 <400> SEQUENCE: 1
23   atggcgcccta tcacggccta tgcccagcag acaaggggcc ttttgggatg cataatcacc 60
24   agcttgaccg gccgggacaa aaaccagggt gaggggtgagg ttcagatcgt gtcaactgct 120
25   gccagactt tcttggcaac ctgcattaac ggggtgtgtt ggactgtcta ccatggagcc 180
26   ggaacaagga ccattgcgtc acctaagggt cctgttatcc agatgtacac caatgtggac 240
27   caagacctcg taggtggcc cgctcccaa ggtgcccgt cattaacacc atgcacttgc 300
28   ggctcctcgg acctttacct ggtcacgagg cacgccgatg tcattcctgt gcgccgacgg 360
29   ggtgatggca ggggcagcct gctttcgccc cggcctatct cttacttgaa aggtcctcgt 420
30   ggaggccctc tgctgtgccc cgcaggacat gccgtaggca tattcagagc cgcggtatgc 480
31   acccgtggag tggctaaggc ggtggacttc atccccgtag agagcttaga gacaaccatg 540
32   aggtccccgg tgttctcaga caactcctcc ccaccagcag tgccccagag ctaccaagtg 600
33   gccacactgc atgctcccac cggcagcggg aagagcacca aggtcccggc cgcatacgca 660
34   gctcagggtt acaagggtgt ggtgctcaac ccctccgttg ctgcaacaat gggcttttgt 720
35   gcttacatgt ccaaggccca tgggattgat cctaacatca ggactggggt gaggacaatt 780
36   actactggca gcccgatcac gtattccacc tacggcaagt tccttgccga cggcgggtgt 840
37   tcaggggggt cttatgacat aataatttgt gacgagtgcc actccacgga tgcaacatcc 900
38   atcttgggca ttggcactgt ccttgaccaa gcagagaccg cgggggcgag actgactgtg 960
39   ctgcgccacc ctaccctcc gggctccgtc actgtgcccc atcctaacat cgaggagggt 1020
40   gctctgtcca ctaccggaga gatccccctt tatggcaagg ctattcccct tgaagcaatt 1080
41   aaggggggga gacatctcat cttctgccac tcaaagaaga agtgcgacga gctcgccgca 1140
42   aaactggtcg cgttgggcgt caatgccgtg gcttactacc gcggccttga tgtgtccgtc 1200
43   atcccgacca gtggtgacgt tgctcgctgt gcaactgacg ccctcatgac cggctttacc 1260
44   ggcgacttcg attcggtgat agactgcaac acgtgtgtca cccagacagt cgacttcagc 1320

```

RAW SEQUENCE LISTING

DATE: 06/08/2005

PATENT APPLICATION: US/09/930,591

TIME: 10:52:17

Input Set : N:\Crf4\Refhold\09_folder\I930591.raw

Output Set: N:\CRF4\06082005\I930591.raw

```

45      cttgacccta ccttcacccat tgagacaatc acgcttcccc aggatgctgt ctcccgtact 1380
46      caacgtcggg gtaggactgg cagaggggaag ccaggcatct acagatttgt ggcaccgggg 1440
47      gagcgctcct ctggcatggt tgactcgtct gtcctctgcg agtgctatga cgcgggttgt 1500
48      gcttggtatg agcttacgcc cgccgagacc acagttaggc tacgagcata catgaacacc 1560
49      ccgggacttc ccgtgtgcc aagaccatctt gaattttggg agggcgctctt tacgggtctc 1620
50      acccacatag acgcccactt cctatcccag acaaagcaga gtggggaaaa ccttccctat 1680
51      ctggtagcgt accaagccac cgtgtgcgct agagctcaag cccctcccc gtcgtgggac 1740
52      cagatgtgga agtgcttgat ccgtctcaag cccaccctcc atggggcaac acctctgcta 1800
53      tatagactgg gcgctgtcca gaatgaagtc accctgacgc acccagtcac caagtatatc 1860
54      atgacatgta tgcggctga cctggaggtc gtcacgagta cctgggtgct cgttggcggc 1920
55      gttctggctg ctttggccgc gtattgccta tccacaggct gcgtgggtcat agtaggtagg 1980
56      attgtcttgt ccggaaagcc ggcaatcata cccgacaggg aagtcctcta ccgggagttc 2040
57      gatgaaatgg aagagtgtctg a                                     2061

```

59 <210> SEQ ID NO: 2

60 <211> LENGTH: 686

61 <212> TYPE: PRT

62 <213> ORGANISM: Artificial Sequence

63 <220> FEATURE:

64 <223> OTHER INFORMATION: Hepatitis C virus NS3/4A peptide

65 <400> SEQUENCE: 2

```

66      Met Ala Pro Ile Thr Ala Tyr Ala Gln Gln Thr Arg Gly Leu Leu Gly
67      1          5          10          15
68      Cys Ile Ile Thr Ser Leu Thr Gly Arg Asp Lys Asn Gln Val Glu Gly
69      20          25          30
70      Glu Val Gln Ile Val Ser Thr Ala Ala Gln Thr Phe Leu Ala Thr Cys
71      35          40          45
72      Ile Asn Gly Val Cys Trp Thr Val Tyr His Gly Ala Gly Thr Arg Thr
73      50          55          60
74      Ile Ala Ser Pro Lys Gly Pro Val Ile Gln Met Tyr Thr Asn Val Asp
75      65          70          75          80
76      Gln Asp Leu Val Gly Trp Pro Ala Pro Gln Gly Ala Arg Ser Leu Thr
77      85          90          95
78      Pro Cys Thr Cys Gly Ser Ser Asp Leu Tyr Leu Val Thr Arg His Ala
79      100         105         110
80      Asp Val Ile Pro Val Arg Arg Arg Gly Asp Gly Arg Gly Ser Leu Leu
81      115         120         125
82      Ser Pro Arg Pro Ile Ser Tyr Leu Lys Gly Ser Ser Gly Gly Pro Leu
83      130         135         140
84      Leu Cys Pro Ala Gly His Ala Val Gly Ile Phe Arg Ala Ala Val Cys
85      145         150         155         160
86      Thr Arg Gly Val Ala Lys Ala Val Asp Phe Ile Pro Val Glu Ser Leu
87      165         170         175
88      Glu Thr Thr Met Arg Ser Pro Val Phe Ser Asp Asn Ser Ser Pro Pro
89      180         185         190
90      Ala Val Pro Gln Ser Tyr Gln Val Ala His Leu His Ala Pro Thr Gly
91      195         200         205
92      Ser Gly Lys Ser Thr Lys Val Pro Ala Ala Tyr Ala Ala Gln Gly Tyr
93      210         215         220
94      Lys Val Leu Val Leu Asn Pro Ser Val Ala Ala Thr Met Gly Phe Gly

```

RAW SEQUENCE LISTING

DATE: 06/08/2005

PATENT APPLICATION: US/09/930,591

TIME: 10:52:17

Input Set : N:\Crf4\Refhold\09_folder\I930591.raw

Output Set: N:\CRF4\06082005\I930591.raw

```

95      225      230      235      240
96      Ala Tyr Met Ser Lys Ala His Gly Ile Asp Pro Asn Ile Arg Thr Gly
97              245              250              255
98      Val Arg Thr Ile Thr Thr Gly Ser Pro Ile Thr Tyr Ser Thr Tyr Gly
99              260              265              270
100     Lys Phe Leu Ala Asp Gly Gly Cys Ser Gly Gly Ala Tyr Asp Ile Ile
101              275              280              285
102     Ile Cys Asp Glu Cys His Ser Thr Asp Ala Thr Ser Ile Leu Gly Ile
103              290              295              300
104     Gly Thr Val Leu Asp Gln Ala Glu Thr Ala Gly Ala Arg Leu Thr Val
105     305              310              315              320
106     Leu Ala Thr Ala Thr Pro Pro Gly Ser Val Thr Val Pro His Pro Asn
107              325              330              335
108     Ile Glu Glu Val Ala Leu Ser Thr Thr Gly Glu Ile Pro Phe Tyr Gly
109              340              345              350
110     Lys Ala Ile Pro Leu Glu Ala Ile Lys Gly Gly Arg His Leu Ile Phe
111              355              360              365
112     Cys His Ser Lys Lys Lys Cys Asp Glu Leu Ala Ala Lys Leu Val Ala
113     370              375              380
114     Leu Gly Val Asn Ala Val Ala Tyr Tyr Arg Gly Leu Asp Val Ser Val
115     385              390              395              400
116     Ile Pro Thr Ser Gly Asp Val Val Val Val Ala Thr Asp Ala Leu Met
117              405              410              415
118     Thr Gly Phe Thr Gly Asp Phe Asp Ser Val Ile Asp Cys Asn Thr Cys
119              420              425              430
120     Val Thr Gln Thr Val Asp Phe Ser Leu Asp Pro Thr Phe Thr Ile Glu
121              435              440              445
122     Thr Ile Thr Leu Pro Gln Asp Ala Val Ser Arg Thr Gln Arg Arg Gly
123     450              455              460
124     Arg Thr Gly Arg Gly Lys Pro Gly Ile Tyr Arg Phe Val Ala Pro Gly
125     465              470              475              480
126     Glu Arg Pro Ser Gly Met Phe Asp Ser Ser Val Leu Cys Glu Cys Tyr
127              485              490              495
128     Asp Ala Gly Cys Ala Trp Tyr Glu Leu Thr Pro Ala Glu Thr Thr Val
129              500              505              510
130     Arg Leu Arg Ala Tyr Met Asn Thr Pro Gly Leu Pro Val Cys Gln Asp
131              515              520              525
132     His Leu Glu Phe Trp Glu Gly Val Phe Thr Gly Leu Thr His Ile Asp
133     530              535              540
134     Ala His Phe Leu Ser Gln Thr Lys Gln Ser Gly Glu Asn Leu Pro Tyr
135     545              550              555              560
136     Leu Val Ala Tyr Gln Ala Thr Val Cys Ala Arg Ala Gln Ala Pro Pro
137              565              570              575
138     Pro Ser Trp Asp Gln Met Trp Lys Cys Leu Ile Arg Leu Lys Pro Thr
139              580              585              590
140     Leu His Gly Pro Thr Pro Leu Leu Tyr Arg Leu Gly Ala Val Gln Asn
141     595              600              605
142     Glu Val Thr Leu Thr His Pro Val Thr Lys Tyr Ile Met Thr Cys Met
143     610              615              620

```

RAW SEQUENCE LISTING

DATE: 06/08/2005

PATENT APPLICATION: US/09/930,591

TIME: 10:52:17

Input Set : N:\Crf4\Refhold\09_folder\I930591.raw

Output Set: N:\CRF4\06082005\I930591.raw

```

144      Ser Ala Asp Leu Glu Val Val Thr Ser Thr Trp Val Leu Val Gly Gly
145      625                      630                      635                      640
146      Val Leu Ala Ala Leu Ala Ala Tyr Cys Leu Ser Thr Gly Cys Val Val
147                      645                      650                      655
148      Ile Val Gly Arg Ile Val Leu Ser Gly Lys Pro Ala Ile Ile Pro Asp
149                      660                      665                      670
150      Arg Glu Val Leu Tyr Arg Glu Phe Asp Glu Met Glu Glu Cys
151                      675                      680                      685
153 <210> SEQ ID NO: 3
154 <211> LENGTH: 686
155 <212> TYPE: PRT
156 <213> ORGANISM: Artificial Sequence
157 <220> FEATURE:
158 <223> OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A
159 <400> SEQUENCE: 3
160      Met Ala Pro Ile Thr Ala Tyr Ala Gln Gln Thr Arg Gly Leu Leu Gly
161      1                      5                      10                      15
162      Cys Ile Ile Thr Ser Leu Thr Gly Arg Asp Lys Asn Gln Val Glu Gly
163                      20                      25                      30
164      Glu Val Gln Ile Val Ser Thr Ala Ala Gln Thr Phe Leu Ala Thr Cys
165                      35                      40                      45
166      Ile Asn Gly Val Cys Trp Thr Val Tyr His Gly Ala Gly Thr Arg Thr
167      50                      55                      60
168      Ile Ala Ser Pro Lys Gly Pro Val Ile Gln Met Tyr Thr Asn Val Asp
169      65                      70                      75                      80
170      Gln Asp Leu Val Gly Trp Pro Ala Pro Gln Gly Ala Arg Ser Leu Thr
171                      85                      90                      95
172      Pro Cys Thr Cys Gly Ser Ser Asp Leu Tyr Leu Val Thr Arg His Ala
173                      100                     105                     110
174      Asp Val Ile Pro Val Arg Arg Arg Gly Asp Gly Arg Gly Ser Leu Leu
175                      115                     120                     125
176      Ser Pro Arg Pro Ile Ser Tyr Leu Lys Gly Ser Ser Gly Gly Pro Leu
177      130                     135                     140
178      Leu Cys Pro Ala Gly His Ala Val Gly Ile Phe Arg Ala Ala Val Cys
179      145                     150                     155                     160
180      Thr Arg Gly Val Ala Lys Ala Val Asp Phe Ile Pro Val Glu Ser Leu
181                      165                      170                      175
182      Glu Thr Thr Met Arg Ser Pro Val Phe Ser Asp Asn Ser Ser Pro Pro
183                      180                      185                      190
184      Ala Val Pro Gln Ser Tyr Gln Val Ala His Leu His Ala Pro Thr Gly
185                      195                      200                      205
186      Ser Gly Lys Ser Thr Lys Val Pro Ala Ala Tyr Ala Ala Gln Gly Tyr
187      210                     215                     220
188      Lys Val Leu Val Leu Asn Pro Ser Val Ala Ala Thr Met Gly Phe Gly
189      225                     230                     235                     240
190      Ala Tyr Met Ser Lys Ala His Gly Ile Asp Pro Asn Ile Arg Thr Gly
191                      245                      250                      255
192      Val Arg Thr Ile Thr Thr Gly Ser Pro Ile Thr Tyr Ser Thr Tyr Gly
193                      260                      265                      270

```

RAW SEQUENCE LISTING

DATE: 06/08/2005

PATENT APPLICATION: US/09/930,591

TIME: 10:52:17

Input Set : N:\Crf4\Refhold\09_folder\I930591.raw

Output Set: N:\CRF4\06082005\I930591.raw

```

194   Lys Phe Leu Ala Asp Gly Gly Cys Ser Gly Gly Ala Tyr Asp Ile Ile
195           275                      280                      285
196   Ile Cys Asp Glu Cys His Ser Thr Asp Ala Thr Ser Ile Leu Gly Ile
197           290                      295                      300
198   Gly Thr Val Leu Asp Gln Ala Glu Thr Ala Gly Ala Arg Leu Thr Val
199           305                      310                      315                      320
200   Leu Ala Thr Ala Thr Pro Pro Gly Ser Val Thr Val Pro His Pro Asn
201           325                      330                      335
202   Ile Glu Glu Val Ala Leu Ser Thr Thr Gly Glu Ile Pro Phe Tyr Gly
203           340                      345                      350
204   Lys Ala Ile Pro Leu Glu Ala Ile Lys Gly Gly Arg His Leu Ile Phe
205           355                      360                      365
206   Cys His Ser Lys Lys Lys Cys Asp Glu Leu Ala Ala Lys Leu Val Ala
207           370                      375                      380
208   Leu Gly Val Asn Ala Val Ala Tyr Tyr Arg Gly Leu Asp Val Ser Val
209           385                      390                      395                      400
210   Ile Pro Thr Ser Gly Asp Val Val Val Val Ala Thr Asp Ala Leu Met
211           405                      410                      415
212   Thr Gly Phe Thr Gly Asp Phe Asp Ser Val Ile Asp Cys Asn Thr Cys
213           420                      425                      430
214   Val Thr Gln Thr Val Asp Phe Ser Leu Asp Pro Thr Phe Thr Ile Glu
215           435                      440                      445
216   Thr Ile Thr Leu Pro Gln Asp Ala Val Ser Arg Thr Gln Arg Arg Gly
217           450                      455                      460
218   Arg Thr Gly Arg Gly Lys Pro Gly Ile Tyr Arg Phe Val Ala Pro Gly
219           465                      470                      475                      480
220   Glu Arg Pro Ser Gly Met Phe Asp Ser Ser Val Leu Cys Glu Cys Tyr
221           485                      490                      495
222   Asp Ala Gly Cys Ala Trp Tyr Glu Leu Thr Pro Ala Glu Thr Thr Val
223           500                      505                      510
224   Arg Leu Arg Ala Tyr Met Asn Thr Pro Gly Leu Pro Val Cys Gln Asp
225           515                      520                      525
226   His Leu Glu Phe Trp Glu Gly Val Phe Thr Gly Leu Thr His Ile Asp
227           530                      535                      540
228   Ala His Phe Leu Ser Gln Thr Lys Gln Ser Gly Glu Asn Leu Pro Tyr
229           545                      550                      555                      560
230   Leu Val Ala Tyr Gln Ala Thr Val Cys Ala Arg Ala Gln Ala Pro Pro
231           565                      570                      575
232   Pro Ser Trp Asp Gln Met Trp Lys Cys Leu Ile Arg Leu Lys Pro Thr
233           580                      585                      590
234   Leu His Gly Pro Thr Pro Leu Leu Tyr Arg Leu Gly Ala Val Gln Asn
235           595                      600                      605
236   Glu Val Thr Leu Thr His Pro Val Thr Lys Tyr Ile Met Thr Cys Met
237           610                      615                      620
238   Ser Ala Asp Leu Glu Val Val Thr Gly Thr Trp Val Leu Val Gly Gly
239           625                      630                      635                      640
240   Val Leu Ala Ala Leu Ala Ala Tyr Cys Leu Ser Thr Gly Cys Val Val
241           645                      650                      655
242   Ile Val Gly Arg Ile Val Leu Ser Gly Lys Pro Ala Ile Ile Pro Asp

```

VERIFICATION SUMMARY

DATE: 06/08/2005

PATENT APPLICATION: US/09/930,591

TIME: 10:52:18

Input Set : N:\Crf4\Refhold\09_folder\I930591.raw

Output Set: N:\CRF4\06082005\I930591.raw